

# Improving SA4 level estimates from the Labour Force Survey using administrative data models

Description of new model for producing enhanced regional labour force estimates and plans for ongoing outputs

Released 27/06/2023

## On this page

- [Background](#)
- [Quality limitations of direct survey estimates](#)
- [Method used to model SA4 level labour force estimates](#)
- [Preliminary SA4 level estimates](#)
- [Release plans and future enhancements](#)
- [Feedback](#)

## Background

The monthly [Labour Force Survey \(/statistics/labour/employment-and-unemployment/labour-force-australia/latest-release\)](#) (LFS) provides timely information on the labour market activity of the usually resident civilian population of Australia aged 15 years and over. The LFS is designed primarily to provide national estimates, with the secondary design objective of producing state and territory estimates.

While the LFS is not designed to produce regional estimates, these have been possible as a by-product of the large sample size required to produce high quality national estimates. Estimates for Statistical Area Level 4s (SA4s) and Greater Capital City Statistical Areas are published on a monthly basis in [Labour Force, Detailed \(/statistics/labour/employment-and-unemployment/labour-force-australia-detailed/latest-release#labour-market-regions-sa4-\)](#).

Since the regional estimates (for 87 SA4 level regions) are based on sub-samples of the LFS, some of which have relatively small populations, they are of lower quality compared to those produced at state and territory and national levels. However, regional labour force statistics are in high demand, and are often the only timely and regular regional or community level economic information.

The Australian Government provided funding to the Australian Bureau of Statistics (ABS) to improve the quality of SA4 level estimates through the use of administrative data and a new method. This paper:

- describes the new method that will be used to produce improved SA4 level estimates
- provides illustrative results for SA4s, based on a preliminary version of the model
- provides some early thinking on dissemination plans, including timing
- highlights the next steps in enhancing regional labour force statistics

## Quality limitations of direct survey estimates

As the traditional SA4 level estimates from the LFS are based on smaller sample sizes, they are of a lower level of statistical quality compared to those produced at the national and state and territory levels. Over time, large data fluctuations occur across most of the regional labour force data, particularly in regions with smaller populations. These fluctuations can be partly the result of local events (for example, floods), but are generally due to sampling variability rather than changes in underlying labour market conditions.

While these fluctuations do not result in impacts on national or state and territory level data, they can limit the usefulness of individual SA4 level data in monitoring and understanding short-term change in the labour market.

The current regional labour force estimates have large relative standard errors (RSEs), generally in excess of the RSE cautionary threshold used by the ABS, owing to their smaller sample sizes. The regional estimates with the highest volatility, and hence the largest RSEs, tend to be in areas with lower populations, such as outback portions of many states.

Given the high RSE's associated with the SA4 level direct survey estimates, ABS advice has traditionally been to use caution when using and interpreting these data on a month-to-month basis, and to use a smoothing approach (with a rolling 12-month average being the most simple and commonly used approach).

## Method used to model SA4 level labour force estimates

ABS is planning to use an area level time series model to produce modelled labour force estimates for SA4s. The model uses the direct survey estimates from the LFS in combination with the following administrative data sources:

- JobSeeker and Youth Allowance recipients data from the Department of Social Services (DSS)
- Single Touch Payroll (STP) data from the Australian Taxation Office (ATO)

The model draws on the strengths of more granular and frequently updated administrative data to provide a more stable and reliable series of estimates.

The model is based on a Rao-Yu (1994) area level time series model. This model uses auxiliary administrative data, along with the direct (weighted) survey estimates to separately model unemployment and employment.

- The unemployment model relies primarily on a time series relationship with the JobSeeker and Youth Allowance recipients data from DSS.
- The employment model relies primarily on a relationship with the STP data from the ATO, although it also makes some use of the JobSeeker and Youth Allowance recipients data from DSS.

Estimates of persons not in the labour force (NILF) are derived as the residual of the SA4 population that is neither employed nor unemployed.

## Detailed model methodology

The modelling method has explicitly considered the effects of seasonality and survey sampling.

Residual seasonality was included in the unemployment model because of the different seasonal patterns that occur in unemployment relative to DSS JobSeeker recipients, particularly around the summer holidays when labour market activity is reduced. For employment, no residual seasonality was detected or required.

The models also include a sampling error term, which is modelled such that the temporal correlations (and variances) induced by the rotating panel design of the LFS are captured. By modelling the direct survey estimates and their associated sampling error covariance structure the complex sample design of the LFS is properly accounted for. The models also include random effect terms for both areas and time points. These terms help to account for differences between areas and time points that the auxiliary data and seasonal terms could not capture.

The models borrow strength across time by the inclusion of the temporal random effects and by the modelling of the correlated sampling errors. These features improve the accuracy of the modelled estimates, particularly increasing the stability of movement estimates.

The models assume a roughly constant relationship between LFS variables and the auxiliary data over time. When there are government policy changes and/or unusual events such as COVID-19, these relationships may change significantly. Time series corrections have been applied to account for this, particularly COVID-19 effects on unemployment.

The models have been applied separately for each of the larger states. For smaller states, SA4s from other states have been combined into the models, as deemed conceptually and empirically appropriate. In particular:

- Tasmania also uses Victoria in modelling
- Northern Territory also uses a selection of outback SA4s in South Australia, Western Australia and Queensland

The Australian Capital Territory (ACT) is only comprised of one SA4. Given the estimates from these SA4 models are designed to be additive to the existing national and state and territory level estimates, the ACT data in the modelled SA4 series will continue to be the same as the territory level direct survey estimate.

There will be refinements made to the model and source data over the coming months. These include, but are not limited to:

- moving from preliminary to final versions of the administrative data sources
- calibrating the SA4 estimates to published state and territory level estimates – so that the sum of SA4 estimates additive to published state and territory totals
- modelling employment, unemployment and NILF simultaneously



## Preliminary SA4 level estimates

A set of preliminary modelled estimates have been produced for each SA4 – for Employed persons, Unemployed persons and Persons Not in the Labour Force (NILF), and for the associated rates (Employment-to-Population ratio, Unemployment rate and Participation rate). These are provided in the spreadsheet below, for the following periods:

- Unemployed persons – July 2016 through to September 2022
- Employed persons, NILF, Unemployment rate, Employment-to-Population ratio and Participation rate – January 2020 through to September 2022

Employment (and therefore the residual NILF data and associated rates) are only currently available from January 2020 as the STP data is only available for this time period. Additional time series models will be explored to consider how to extend the employment time series back beyond the start of STP reporting.

## Modelled estimates by labour force status, by SA4

[↓ Download XLSX](#)  
[322.69 KB]

The map of Australia below shows the preliminary modelled unemployment rates for SA4s in September 2022 (based on the 2021 ASGS boundaries).

## Preliminary modelled SA4 Unemployment rate, September 2022 (Original)

Loading map...

Data for the ACT are the direct survey estimates.

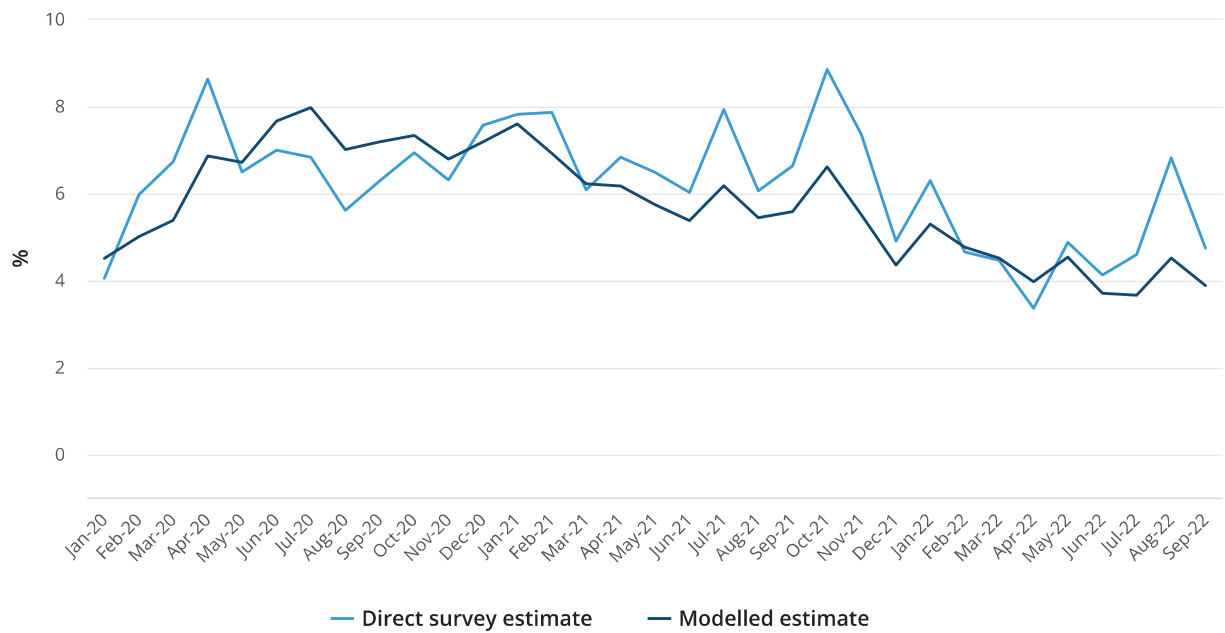
Western Australia - Outback (North) and Western Australia - Outback (South) are modelled as a combined SA4.

The graphs below compare the direct LFS estimates and the preliminary modelled estimates – for Employed persons and the Unemployment rate – for a selected metropolitan and regional area in each state and in the Northern Territory.

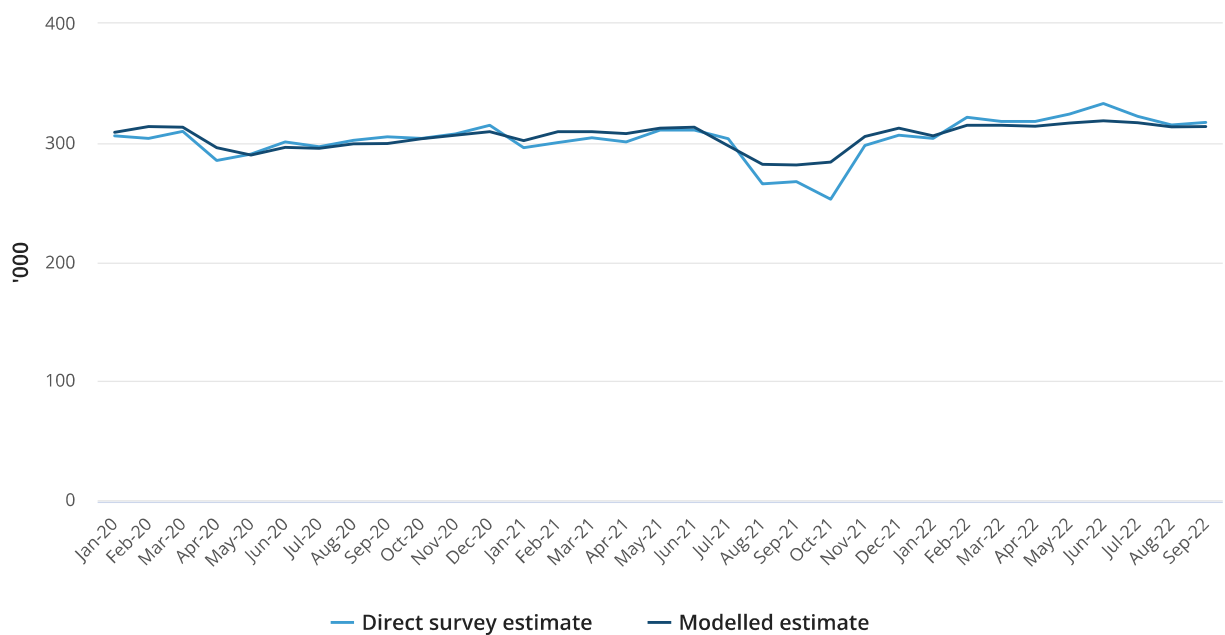
### New South Wales

#### Unemployment rate, Sydney - Inner South West (Original)

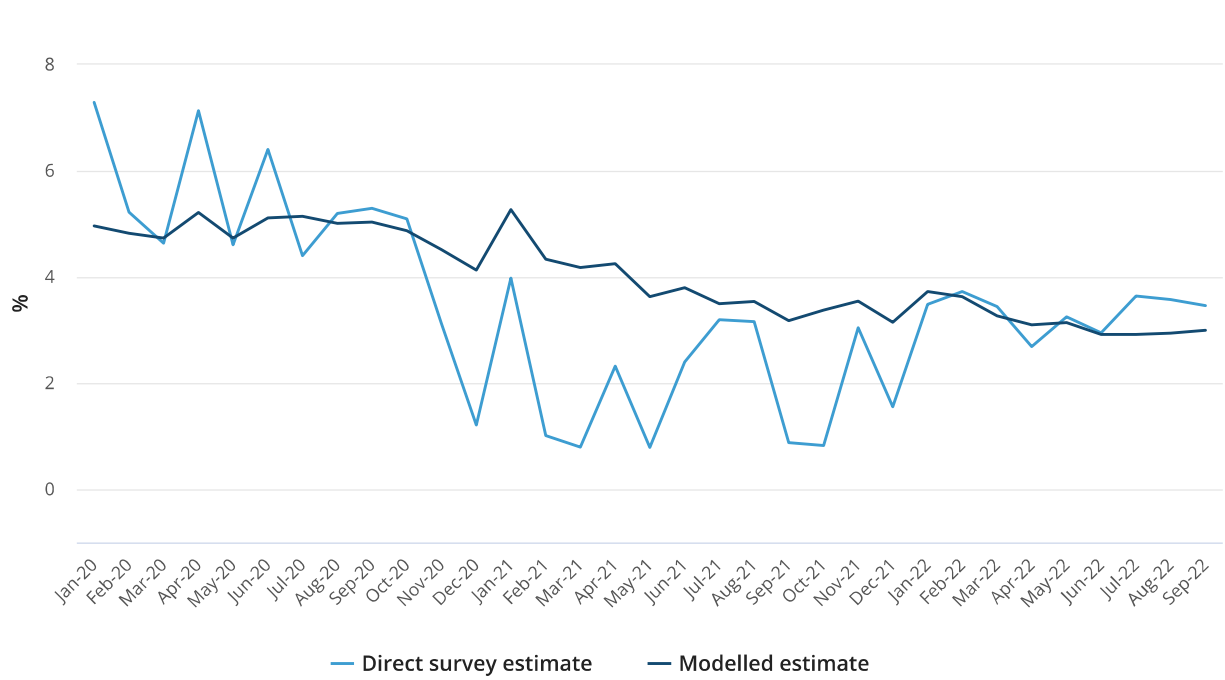
Feedback



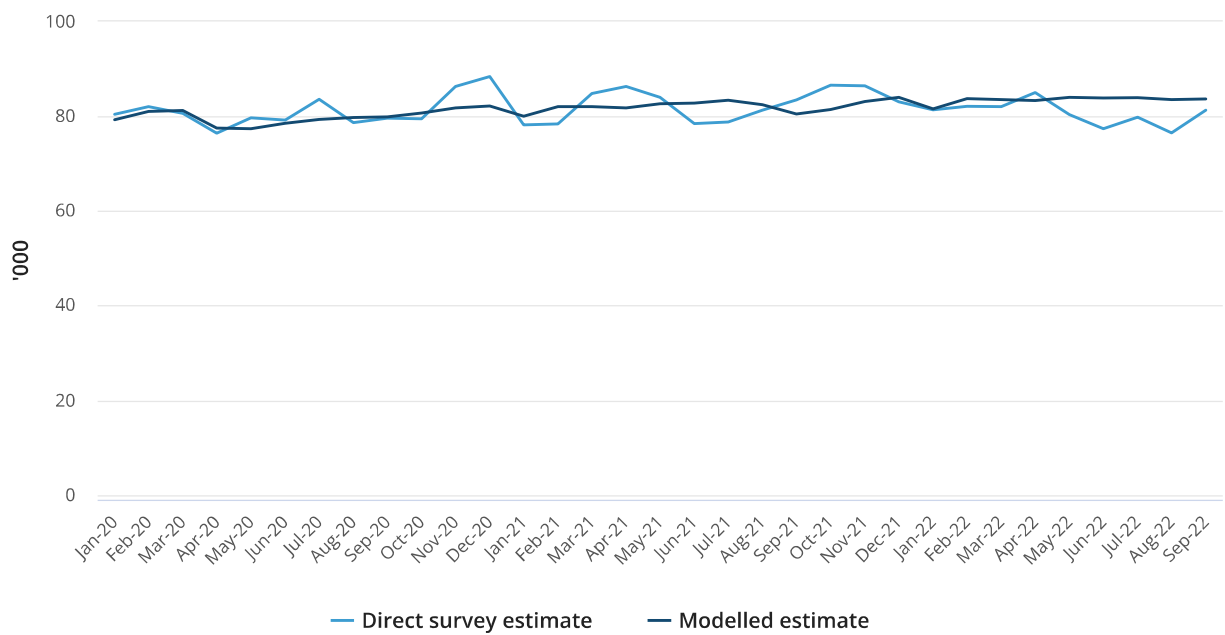
## Employed persons, Sydney - Inner South West (Original)



## Unemployment rate, Riverina (Original)



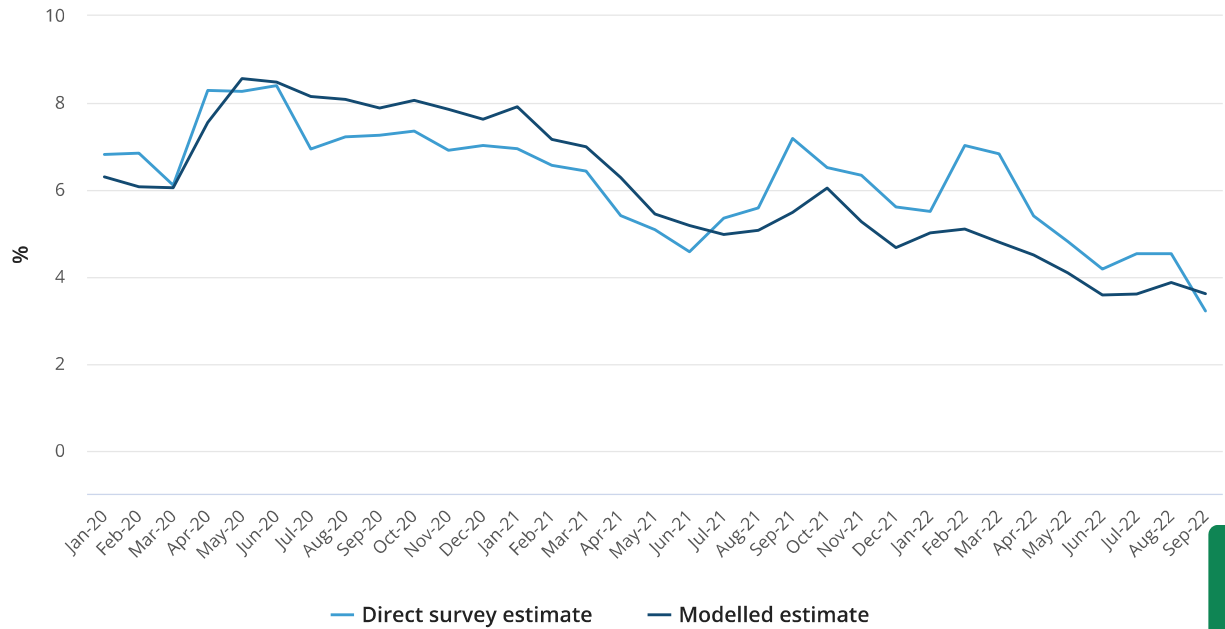
## Employed persons, Riverina (Original)



Feedback

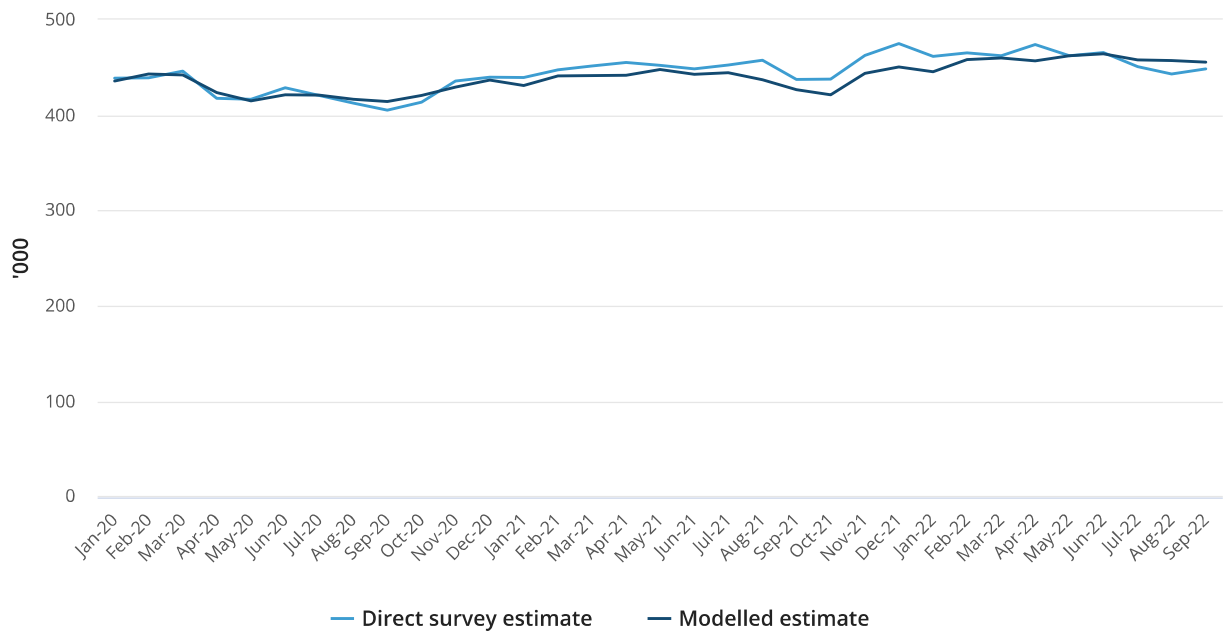
## Victoria

### Unemployment rate, Melbourne - South East (Original)

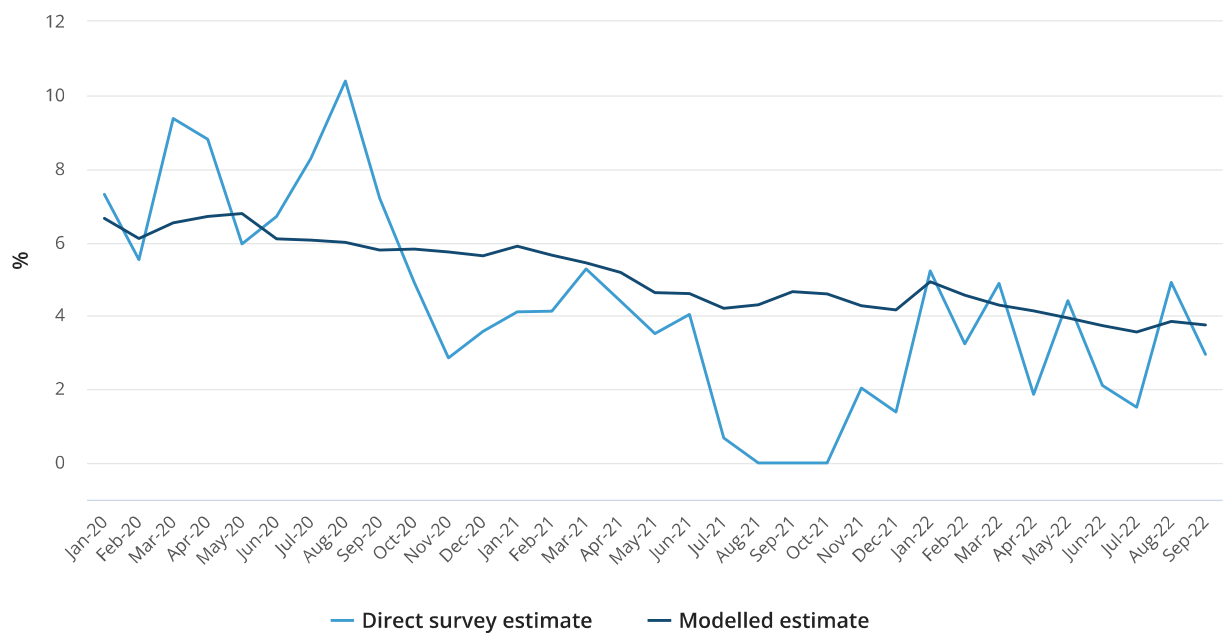


Feedback

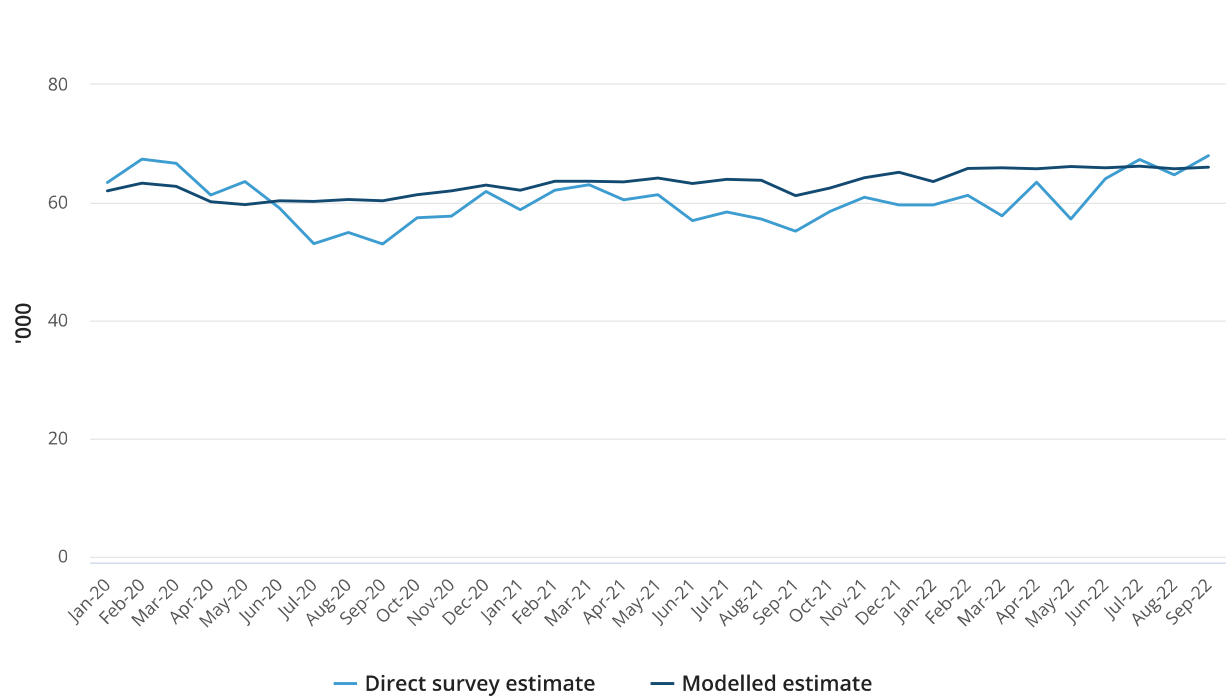
### Employed persons, Melbourne - South East (Original)



## Unemployment rate, Shepparton (Original)

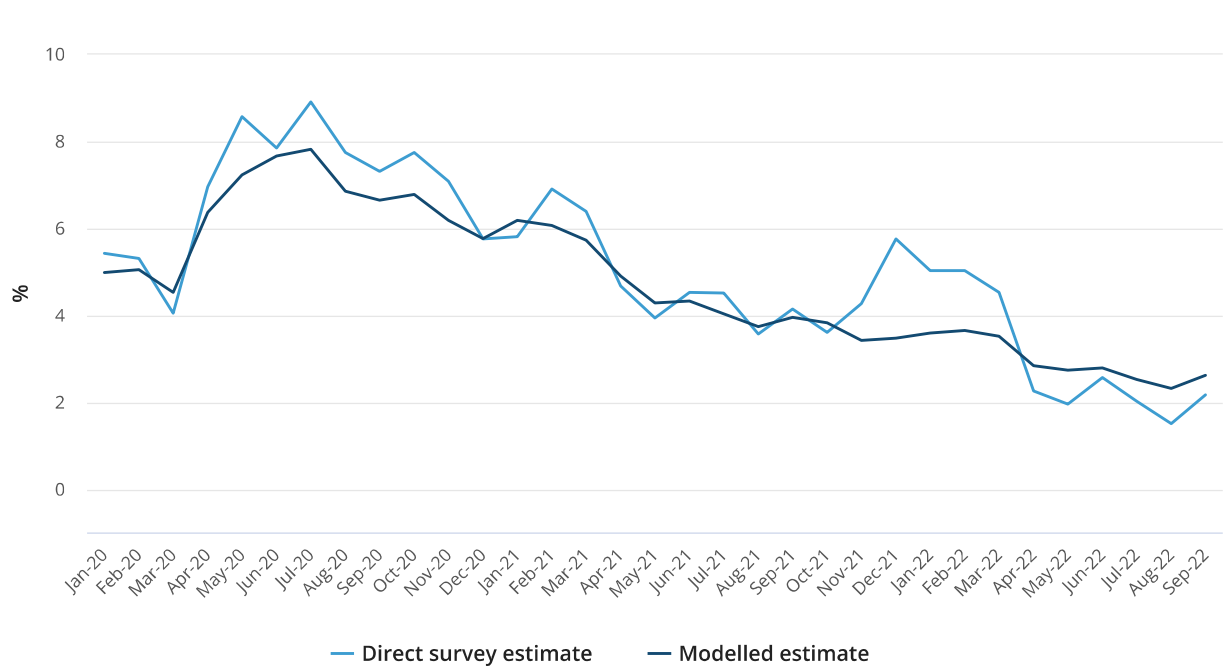


## Employed persons, Shepparton (Original)

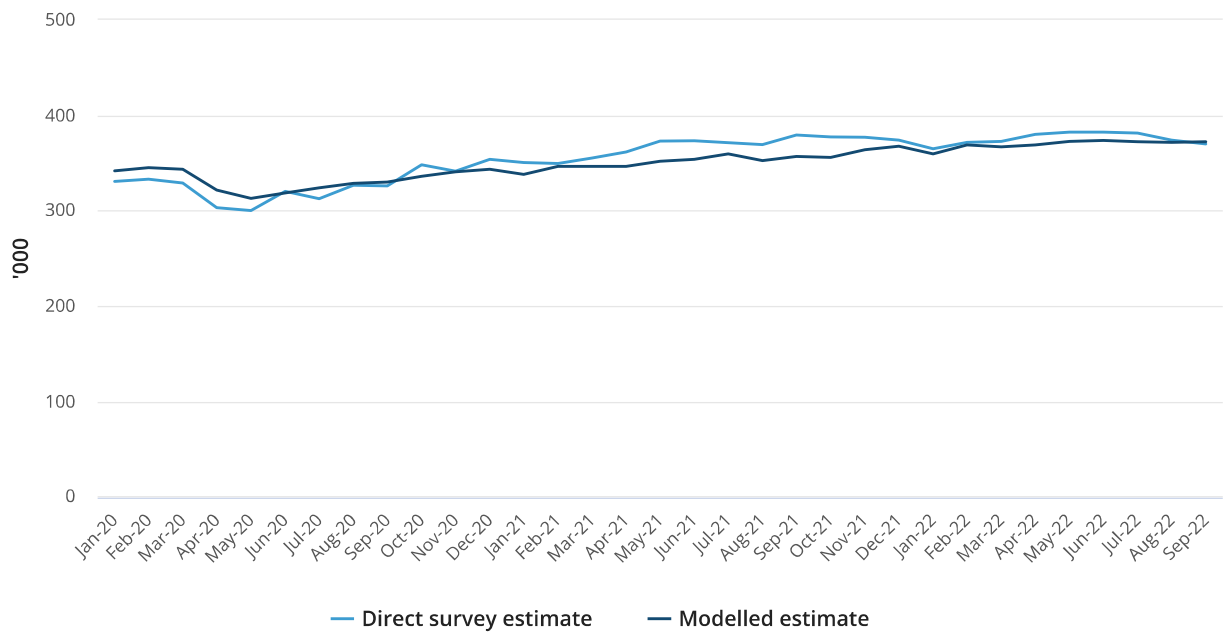


## Queensland

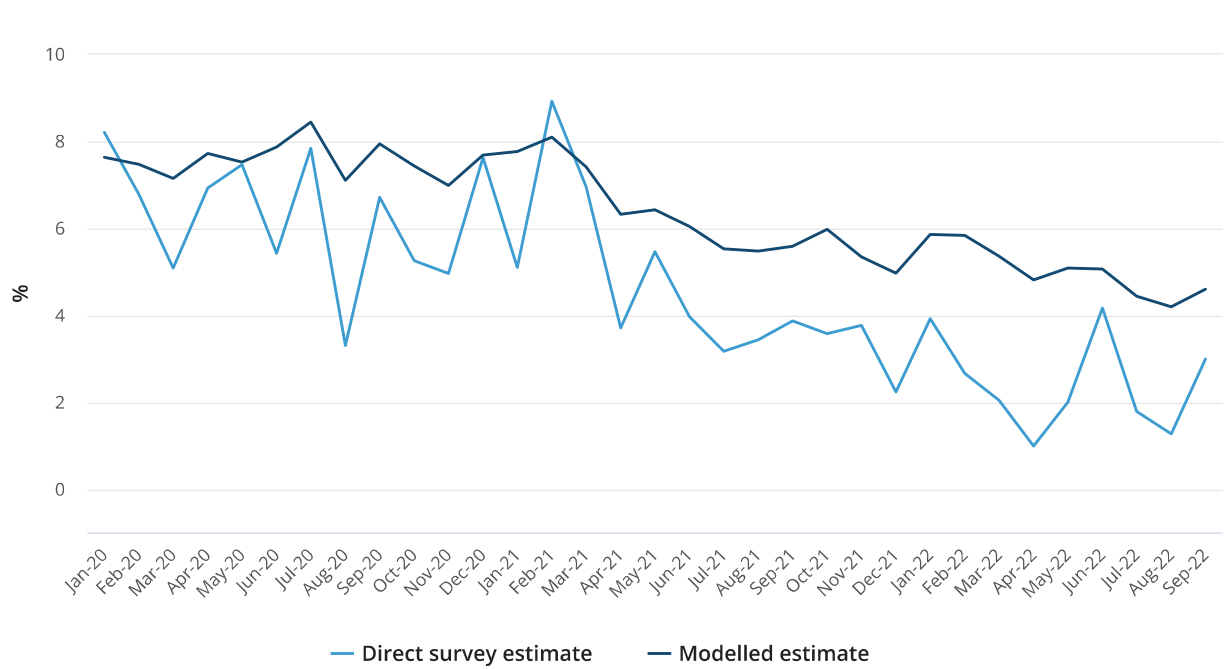
### Unemployment rate, Gold Coast (Original)



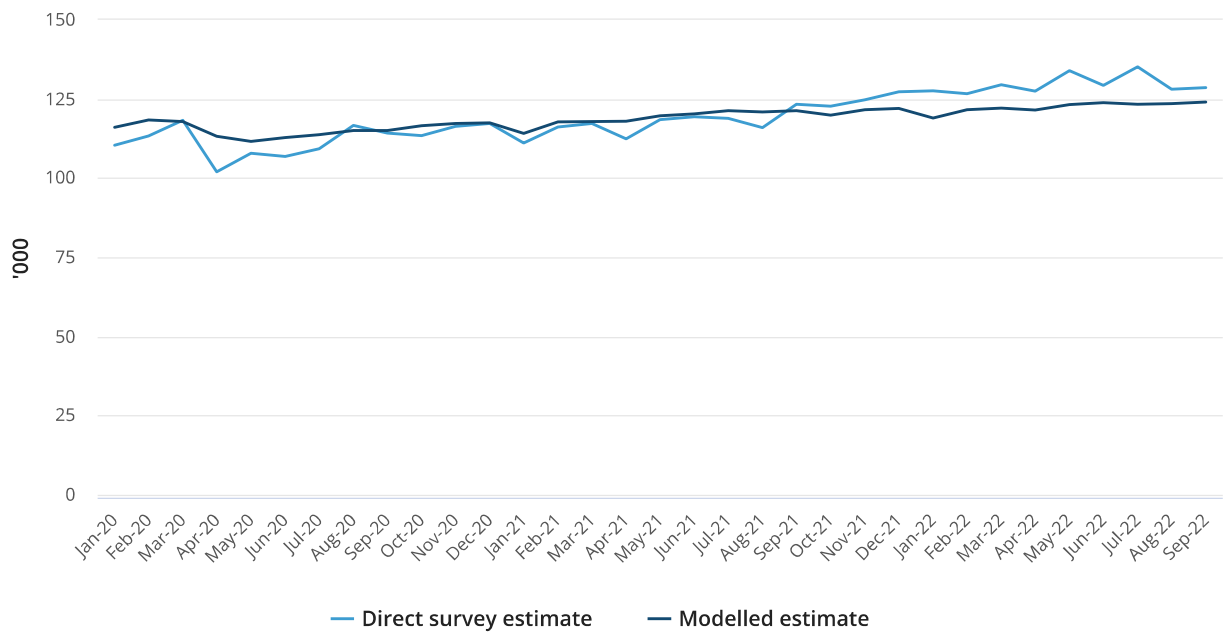
## Employed persons, Gold Coast (Original)



## Unemployment rate, Townsville (Original)



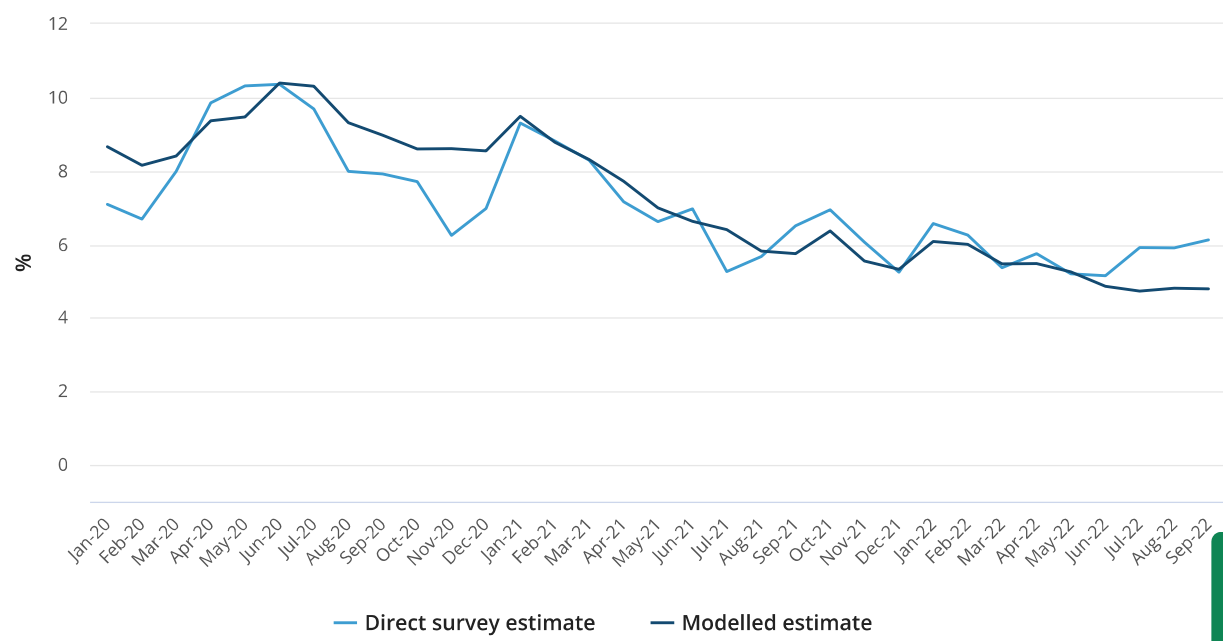
## Employed persons, Townsville (Original)



Feedback

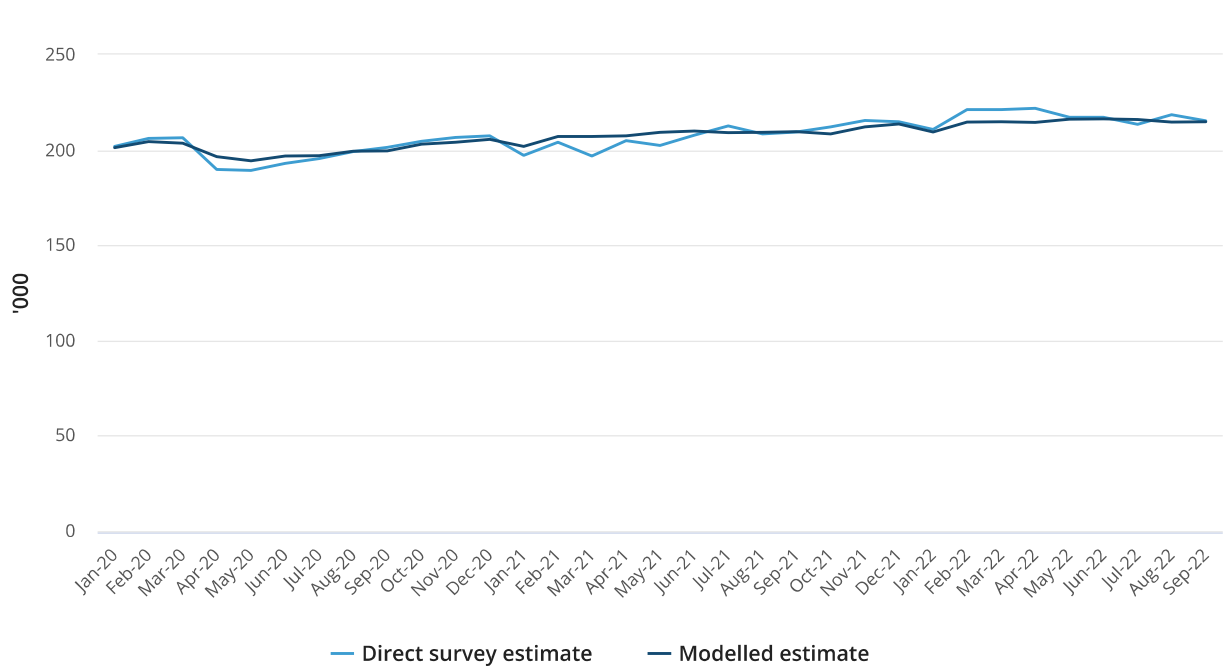
South Australia

Unemployment rate, Adelaide - North (Original)

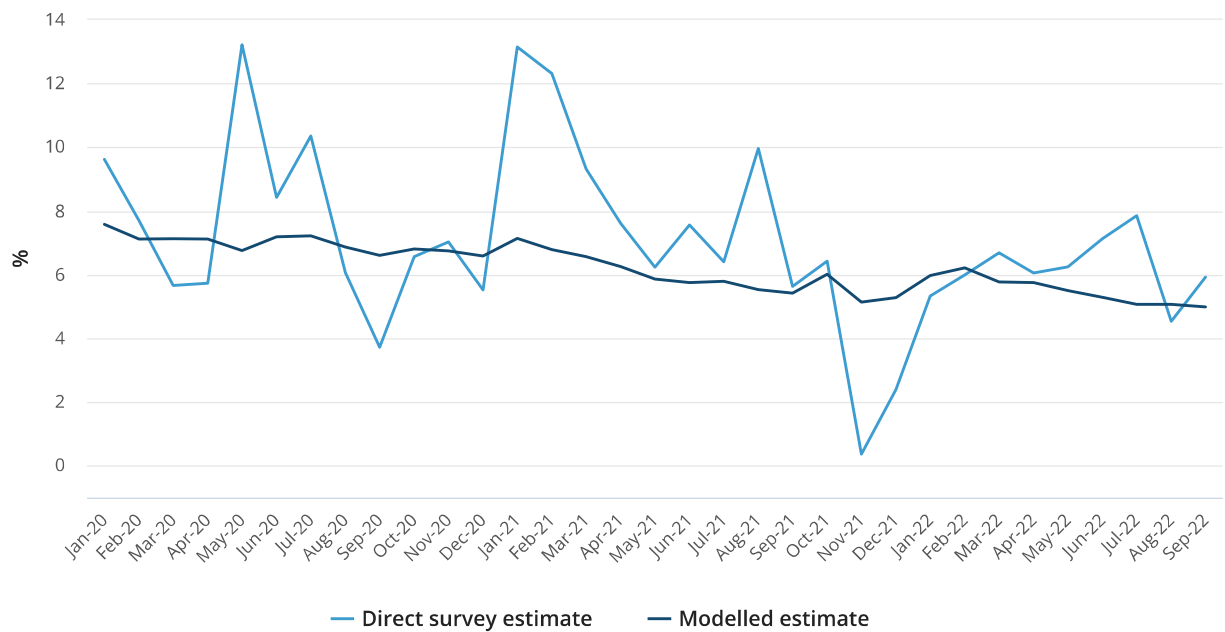


Feedback

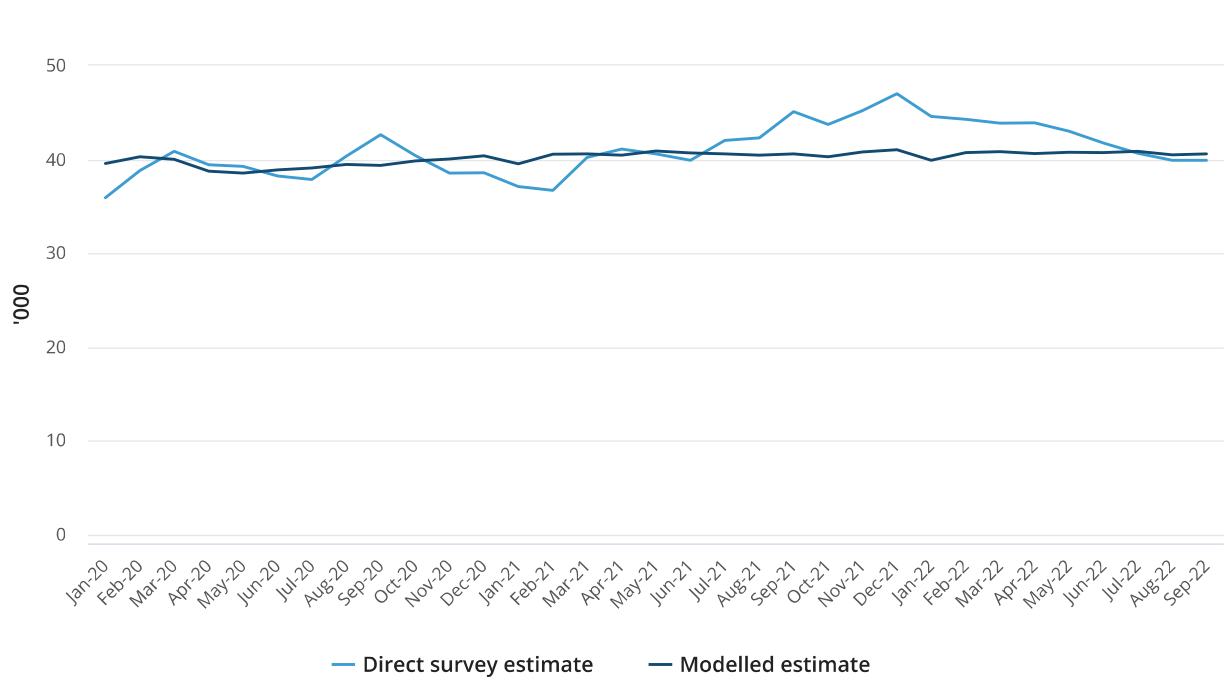
Employed persons, Adelaide - North (Original)



## Unemployment rate, South Australia - Outback (Original)

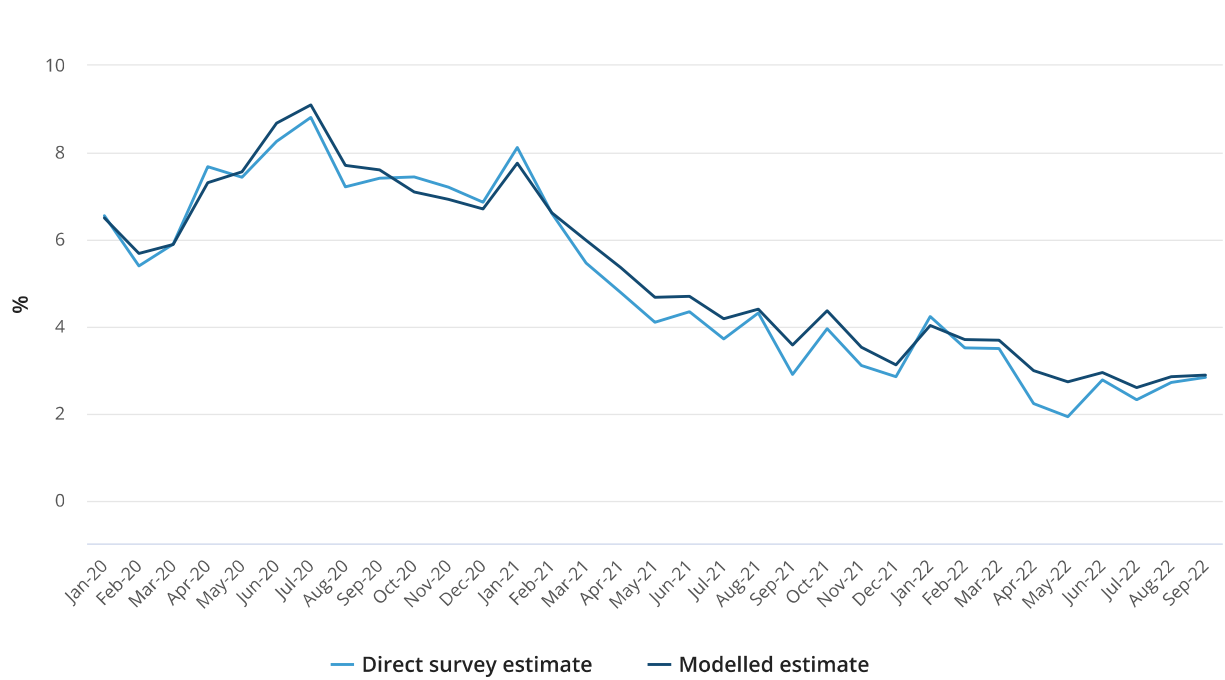


## Employed persons, South Australia - Outback (Original)

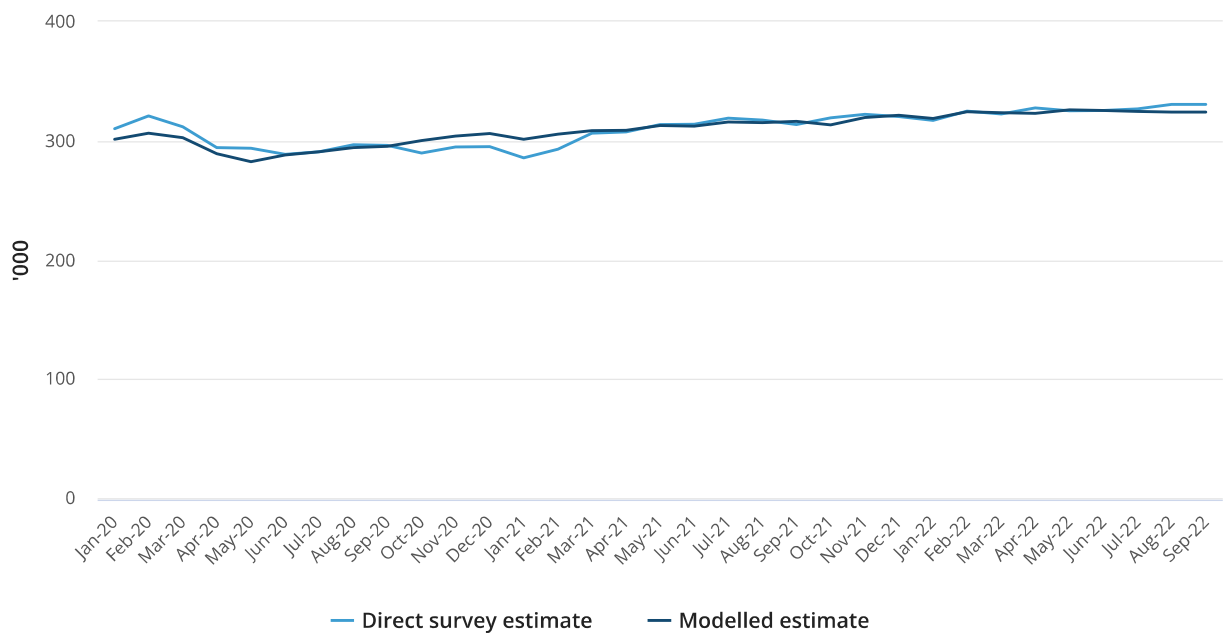


## Western Australia

### Unemployment rate, Perth - North West (Original)

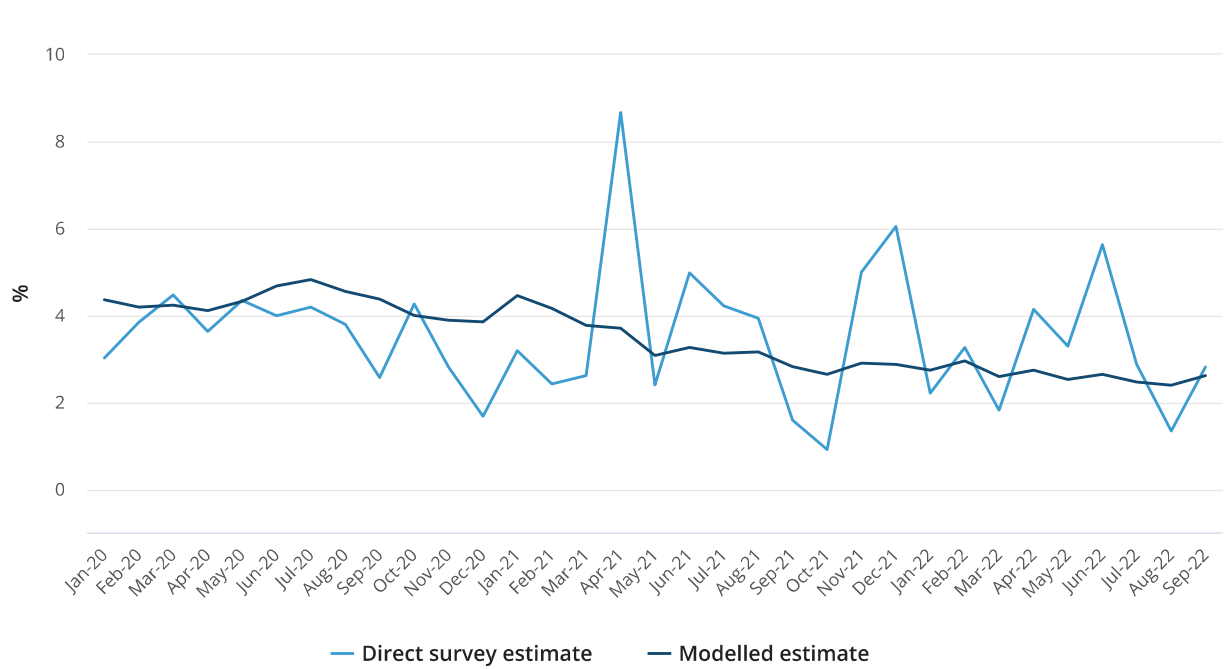


## Employed persons, Perth - North West (Original)



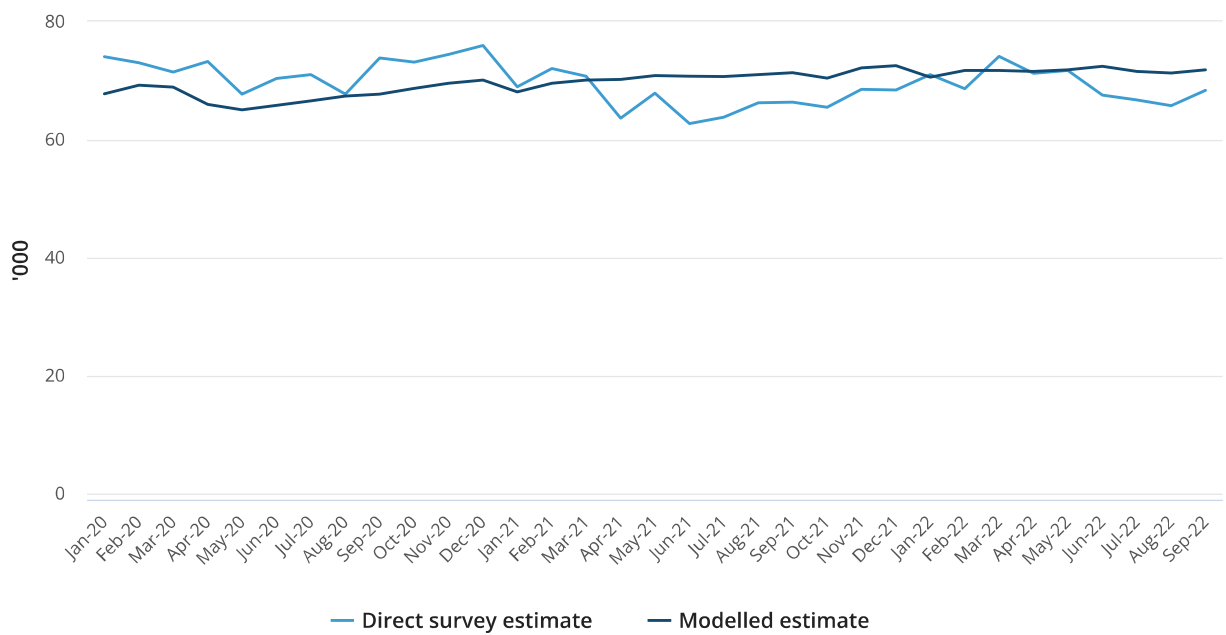
## Unemployment rate, Western Australia - Wheat Belt (Original)

Feedback



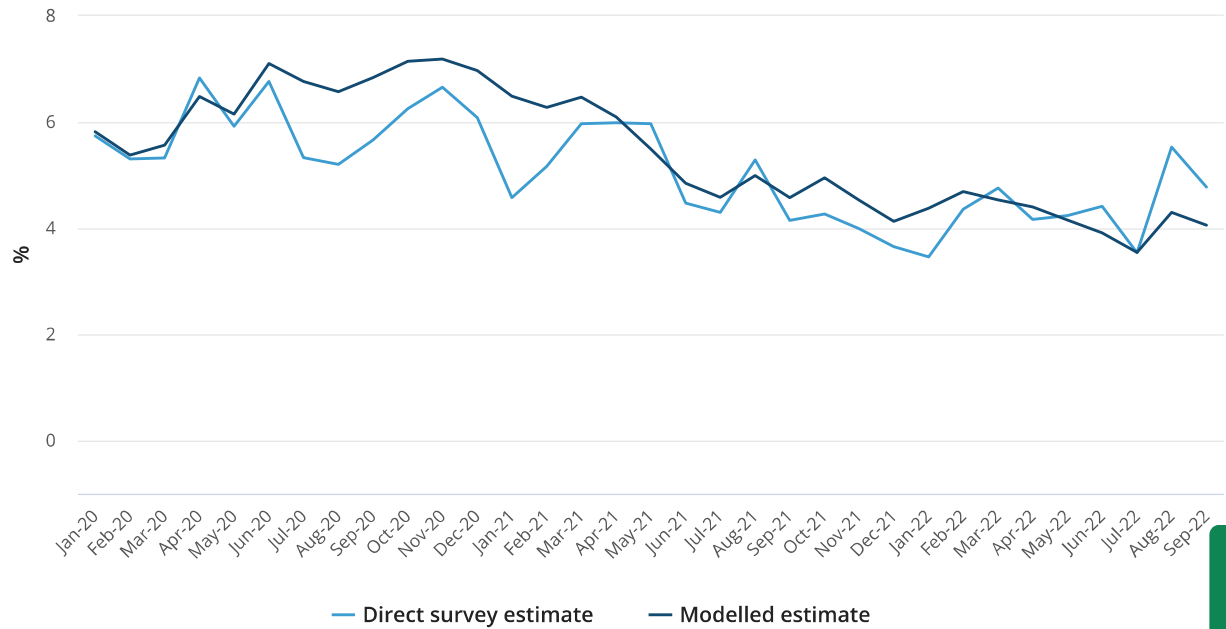
## Employed persons, Western Australia - Wheat Belt (Original)

Feedback



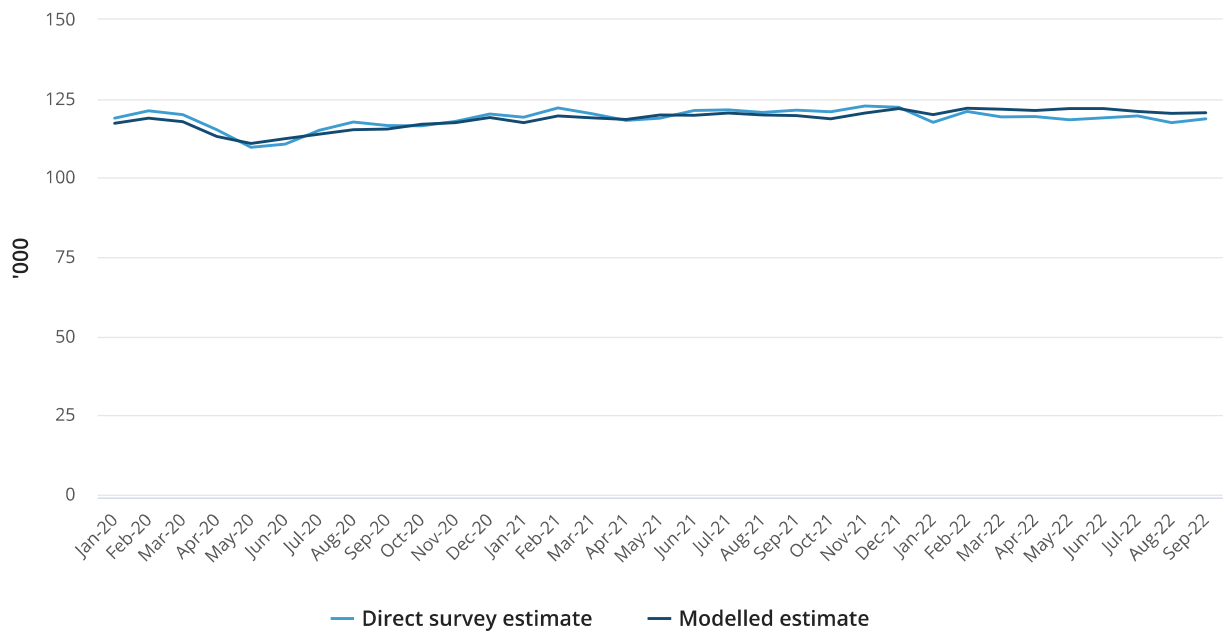
## Tasmania

### Unemployment rate, Hobart (Original)

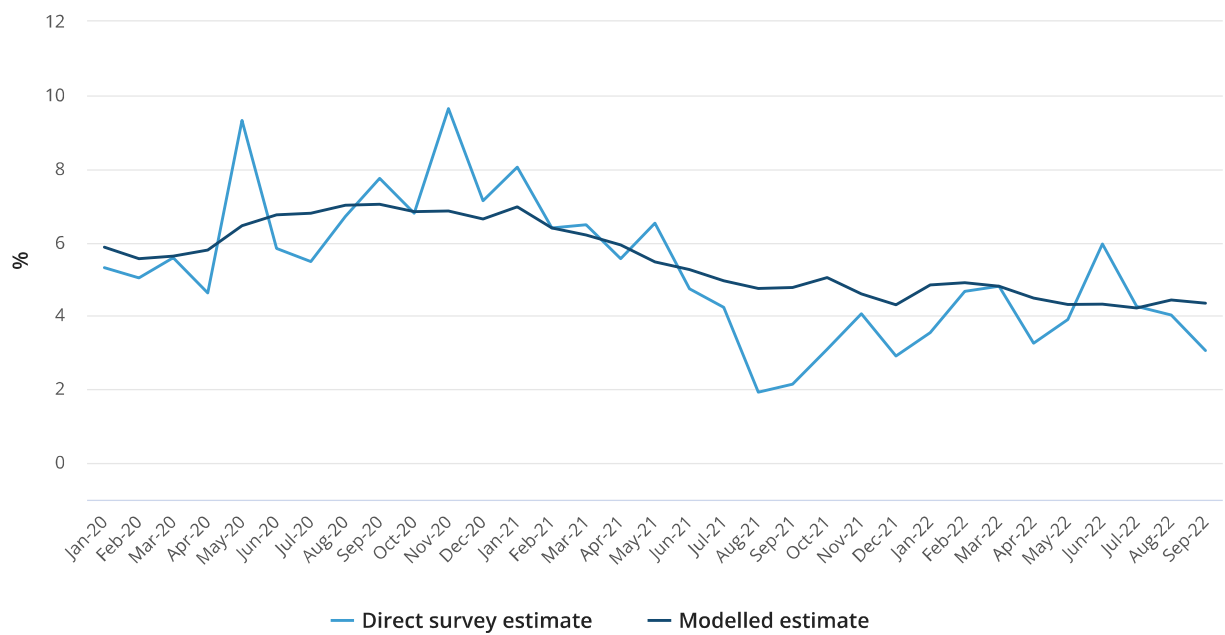


Feedback

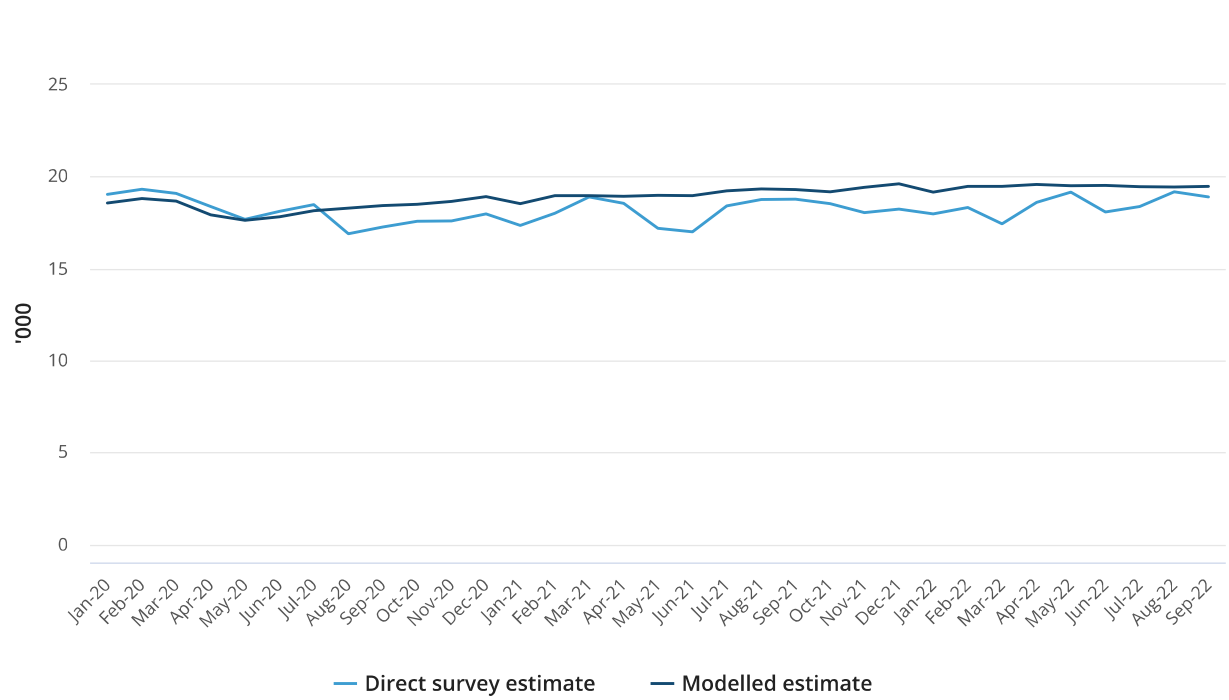
### Employed persons, Hobart (Original)



## Unemployment rate, Tasmania - South East (Original)

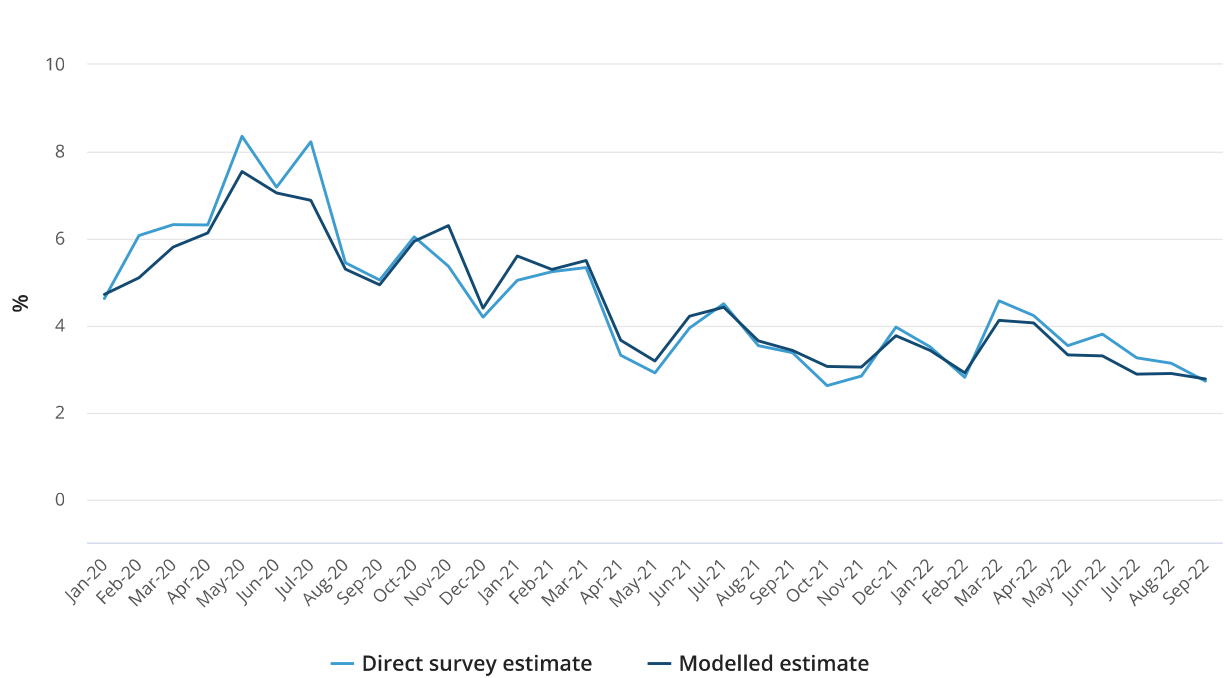


## Employed persons, Tasmania - South East (Original)



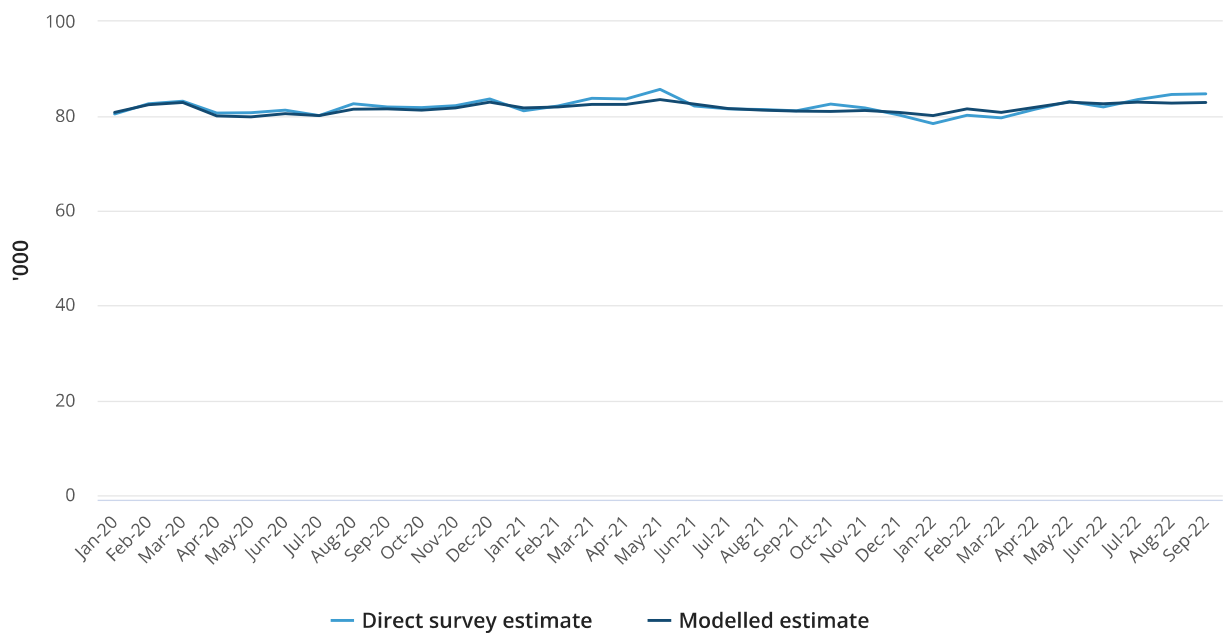
## Northern Territory

### Unemployment rate, Darwin (Original)

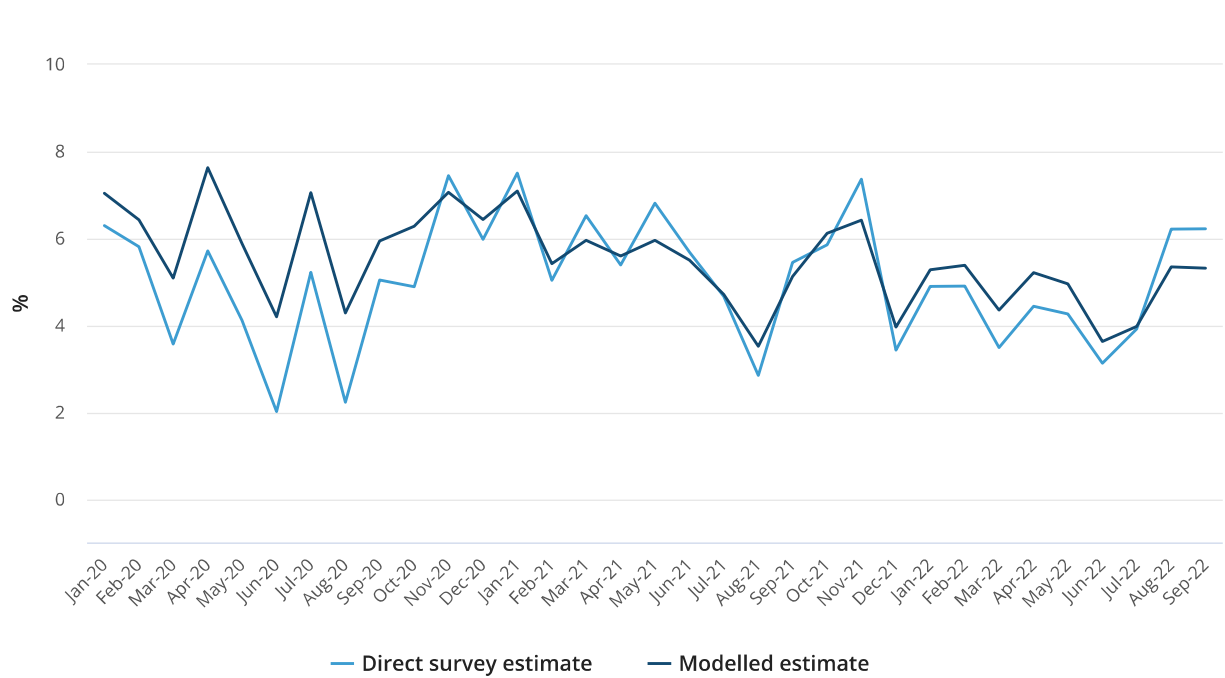


## Employed persons, Darwin (Original)

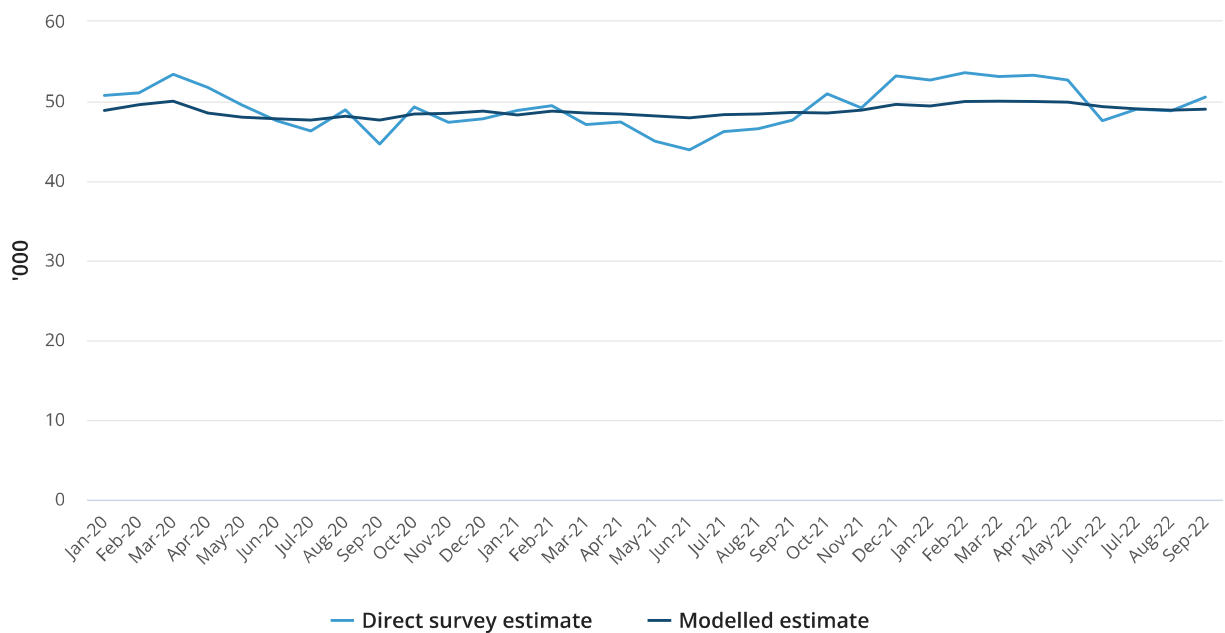
Feedback



## Unemployment rate, Northern Territory - Outback (Original)



## Employed persons, Northern Territory - Outback (Original)



Feedback

## Release plans and future enhancements

The modelled SA4 level labour force estimates will be implemented into the regular set of LFS outputs, and be published on a monthly basis. The first monthly release of the modelled SA4 estimates is expected to be published with the January 2024 LFS, in February 2024.

The modelled estimates will initially be added into [Labour Force, Detailed \(/statistics/labour/employment-and-unemployment/labour-force-australia-detailed/latest-release\)](#) around a week after it is first published. The detailed release comes out a week after the main LFS release, and contains detailed data not included in the first release including the direct survey SA4 estimates. Eventually, the modelled estimates will be published at the time of the detailed release (i.e. without the one week delay), at the same time as the direct survey estimates.

The direct survey estimates also contain more detailed disaggregations (e.g. part-time/full-time, age and sex breakdowns) than is currently produced for the modelled estimates. Further enhancements to the model to enable the production of further disaggregation will be explored.

The current SA4 level direct survey estimates will continue to be published for some time, before eventually being replaced by the modelled estimates. The ABS will provide further updates before making additional changes.

Later this year, and prior to implementing the regular monthly outputs, an article will be released in [Labour Force, Australia \(/statistics/labour/employment-and-unemployment/labour-force-australia/latest-release\)](#) to provide updated estimates (with more up-to-date data, and based on a near-final version of the model) as well as further detail on any final refinements to the model or release plans.

Over the next year the ABS will also be exploring areas for further development of modelled labour force estimates

- additional detail for the SA4 level estimates including a breakdown by full-time/part-time, sex and age groups
  - methods to produce estimates at lower levels of geography - modelled estimates at the SA3, SA2 and LGA level.
- Further information on this work will be published in mid-2024.

## Feedback

If you would like to provide feedback or have any questions, please email [labour.statistics@abs.gov.au](mailto:labour.statistics@abs.gov.au).